CHAPTER III
RESEARCH METHODOLOGY

3.1 Research Method

The pre-experimental is used as a research method in this research. According to Frey (2018) pre-experimental designs are research methods that involve observing a subject or a group after a treatment that has been applied in order to determine if the treatment has the potential to cause change. The dependent variable in this study is students' creativity and concept mastery in heat transfer topic (O). While the learning approach used is STEAM based learning as an independent variable (X).

3.2 Research Design

The research design used in this research is one group pre-test post-test design. In this research design, a single group is measured or observed before and after being treated to some sort of treatment (Fraenkel, Wallen, & Hyun, 2011). These design provide plausible hypotheses explaining a $O_1 - O_2$ difference, as opposed to the hypothesis that X caused the difference, which means they are vulnerable to the possibility that a factor other than the independent variable will cause a difference between the pretest and posttest results (Babbie, 2013; Campbell et al., 1963). In this study, the researcher uses one group pre-test post-test design to observe and measured the effect after the STEAM based learning treatment is implemented. The design is shown in the Table 3.1.

Table 3.1
One Group Pre-test Post-test Design

O_1	X	O_2
Pre-test	Treatment (STEAM based learning)	Post-test

Information:

O1 = Condition before treatment (initial of students' concept mastery)

X = Treatment (Implementation of STEAM based learning)

O2 = Condition after treatment (final of students' concept mastery and creativity)

3.3 Classroom Treatment

The learning activity in this research was carried out across five meeting, each lasting around 40-50 minutes. The pre-test is held concurrently with the first meeting. The rest of the meetings execute STEAM-based learning. Post-test is held in the last meeting. The five stages of STEAM-PjBL proposed by Lou et al. (2017) were employed in this research, including the planning, implementation, presentation, evaluation, and correction stage. The learning activities and each stage of STEAM-PjBL learning can be seen in Table 3.2.

Table 3.2
The Learning Activities and Each Stage of STEAM-PjBL

Meeting	STEAM PjBL Stages	Activity
1st	-	Conducting pre-test by using nearpod application.
2nd	Preparation	At this meeting, the preparation stage is begin. The activities carried out during the preparation stage including stimulating their knowledge, identifying the problems and finding solution, determining materials, and create product design.
3rd	Implementation	During this stage, students are conducted product creation, product testing, and product identification.
4th	Presentation, Evaluation, and Correction	Firstly, implementing presentation stage. Students were asked to present the design ideas and products, communicate issues, and address their problems throughout the implementation stage. Then, applying the evaluation stage, which included activities to evaluate each group product. Lastly, the correction stage, students enhance their products based on feedback obtained during the evaluation stage.
5th	-	Conducting prost-test by using nearpod application.

3.4 Population and Sampling

This research was conducted in one Private Junior High School in Bandung. The school was using Cambridge curriculum in the teaching-learning process. English is one

of the languages that are used as a medium of instruction and communication in a daily basis at the school. The researcher took one class in 8th grades as the sample. None of participants had received STEM–STEAM before. There are 21 students from 8B class as the experimental class which learn the topic about heat transfer, the instruction was carried out with the STEAM approach. Therefore, due to pandemic COVID-19 era, the data collection was done online in April 2021. The sampling technique used convenience sampling. A convenience sample is a group of people who readily and available for study (Fraenkel et al., 2011).

3.5 Operational Definition

An operational definition of this research is stated in order to avoid misconceptions. The terminologies are expressed as followed:

1) Students' Creativity

   Students' creativity in this research is focused on the skill of the students to provide innovation to provide a useful solution as they are able to solve problems. The students are divided into several group to provide collaborative learning. Therefore, students' final work will be examined to measure their creativity. The creativity is the measured by the CPAM rubric which divides the dimension of creativity into three aspects, such as novelty, resolution, and elaboration & synthesis (Besemer & Treffinger, 1981). This rubric will be used to assess students' creativity from their project as the product in learning heat transfer.

2) Students' Concept Mastery

   Students' concept mastery in this research focus on three level of cognitive domain based on Bloom's taxonomy revised version, which include C1 (remembering), C2 (understanding), and C4 (analyzing). This competence was measured by objective test which consisted 15 multiple choice questions to test students' understanding about the concept of heat transfer. The objective test is distributed before and after implementing the treatment, thus the improvement of students' concept mastery is being recognized.

3) STEAM Based Learning

STEAM based learning refers to the approach that is implemented by integrating it with project based learning (PjBL). It conduct five stages: preparation; implementation; presentation; evaluation; and correction (Lou et al., 2017). Each stage will encourage students to be actively involved in completing a project to solve the problems.

3.6 Assumption

According to the literature review, it could be assumed that:

1) STEAM based learning would become a useful learning approach that ease students to master the concept and get better score result in final exam.
2) STEAM based learning activities would encourage the development of varied abilities as well as the stimulation of creativity in project-based practice.

3.7 Hypothesis

To analyze hypothesis more significant by using pretest and posttest, the hypothesis for this research is:

1) H0: There is no difference on students' concept mastery after implementing STEAM based learning in heat transfer topic.
2) H1: There is a difference on students' concept mastery after implementing STEAM based learning in heat transfer topic.

3.8 Research Instrument

In this research, research instrument is very important to be used for gaining data. There were three types of instruments that were used in this research. Both instruments are explained below:

1) Objective Test

Objective test based on Bloom's Revised version was used to measure students' concept mastery before and after implementing STEAM based learning (Wandari et al., 2018). It consists of Pre-test and Post-test. A pre-test is used to determine prior knowledge, and a post-test is used to determine if cognitive mastery is improving or not. The objective test that are used in this objective test is 15 question multiple choices, it consisted of C1 (remembering), C2 (understanding), and C4 (analyzing). Before being utilize as a research instrument, the test items were judged by multiple experts in the

subject and tested on students. After going through the judging procedure, the results were either approved, changed, or removed. Table 3.3 shows the test item.

Table 3.3
Objective Test of Heat Transfer Topic

Subtopic	C1	C2	C4	Total
Heat Transfer	10			1
Conduction		2, 14, 7	4	4
Convection		3, 11		2
Radiation	1	8, 12, 13		4
Evaporation	6	5	9	3
Flask	15			1

a. Validity

Validity has been described as the appropriateness, correctness, meaningfulness, and usefulness of the specific conclusions researchers draw based on the data they gather (Fraenkel et al., 2011). It is a descriptive term used to describe a measure that correctly represents the idea that it is designed to assess. In this research, validity was used to determine whether the instrument was capable of measuring students' concept mastery or not. The researcher utilized a software named ANATES to determine the validity value. The interpretation of the validity value is shown in Table 3.4.

Table 3.4
The Interpretation of Validity Value

The amount of r value	Interpretation
$0{,}80 < r_{xy} \leq 1{,}00$	Very High
$0{,}60 < r_{xy} \leq 0{,}79$	High
$0{,}40 < r_{xy} \leq 0{,}59$	Enough
$0{,}20 < r_{xy} \leq 0{,}39$	Low
$0{,}00 < r_{xy} \leq 0{,}19$	Very Low

(Arif, 2014)

b. Reliability

Reliability is referred to the quality of a measuring procedure that suggests the same data that has been gathered each time in repeated observations of the same occurrence (Consistency) (Fraenkel et al., 2011). The interpretation of the reliability value is shown in Table 3.5.

Table 3.5
The Interpretation Value of Reliability

Gained r value	Interpretation
0,80 – 1,00	Very High
0,60 – 0,79	High
0,40 – 0,59	Enough
0,20 – 0,39	Low
0,00 – 0,19	Very Low

(Arif, 2014; Dhamayanti et al., 2017)

c. Difficulty level

The degree of difficulty that students have in answering an item is referred to the difficulty level. The difficulty index level is calculated by dividing the number of students who are able to answer the question correctly with the total number of students participated in the test (Marie & Edannur, 2015). The greater the value for the difficulty index, the easier the things; conversely, as the number for the difficulty index decreases, the items get more complex (Arif, 2014; Johari et al., 2011). According to Mok (1995) (as cited in Johari et al., 2011) stated that the difficulty index, which categorizes test questions into three degrees of difficulty: easy, moderate, and difficult, has been recognized as one method for assessing the difficulty level in test questions. Furthermore, the item difficulty is defined as the percentage of students that correctly answered the item, which is usually referred to as the p-value. (Marie & Edannur, 2015). The interpretation value of difficulty level is shown in Table 3.6.

Table 3.6
Interpretation Value of Difficulty Level

Difficulty Index	Interpretation
$P < 0.30$	Difficult
$0.30 \leq P \leq 0{,}70$	Moderate
$P > 0.70$	Easy

(Arif, 2014; Johari et al., 2011)

d. Discriminating power

Discriminating power was utilized in a test item to differentiate between strong or high-scoring students and weak or low-scoring students. According to Marie et al. (2015) stated that discriminating power assesses how well a particular item

discriminates between examinees in the function or ability tested by the item. This number is between 0.00 and 1.00. The higher the value, the greater the item's discrimination. A highly discriminating item means that students with high test scores got the item right whereas students with low test scores got it wrong (Marie & Edannur, 2015). The interpretation value of discriminating power is shown in Table 3.7.

Table 3.7
Discriminating Power

The amount of r value	Interpretation
$0{,}00 < D \leq 0{,}20$	Poor
$0{,}20 < D \leq 0{,}40$	Satisfactory
$0{,}40 < D \leq 0{,}70$	Good
$0{,}70 < D \leq 1{,}00$	Excellent
D = Negative	Question is deleted

(Arif, 2014; Marie & Edannur, 2015)

e. Distractor

The wrong options, often known as distractors, are purposefully included among the response options to draw students who do not know the proper answer to a test question (Haladyna & Downing, 1989). Writing high-quality distractors is an important stage in the design of items and tests. Given the item content and the amount of possible distractors, multiple-choice questions should ideally present as many options as feasible (Fulcher & Davidson, 2013; Tarrant, Ware, & Mohammed, 2009). But the most important thing about distractor that it's not only develop the *number* of distractors, but its *quality* (Haladyna & Downing, 1989).

Participants who are found to have inadequate knowledge or experience with the concept are anticipated to select the distractors, whereas those with adequate knowledge and experience will select the proper answer (Fulcher & Davidson, 2013; Tarrant et al., 2009). When people with the necessary knowledge and expertise are unable to distinguish between distractors and the correct response, the question may need to be changed (Boateng, Neilands, Frongillo, Melgar-Quiñonez, & Young, 2018). Non-functional distractors must

be eliminated and replaced with effective ones (Haladyna & Downing, 1989). To find out how many distractors there are in every question, measure the number of students who select the wrong answer.

2) Research Instrument Development and Analysis

The objective test utilized in this research was multiple choice, and it needed to be validated before being utilized as a research instrument. The test item must be validated to verify its validity, reliability, difficulty level, and discriminating power. The objective test consists of 40 questions before judgment by experts. It is used to measure students' concept mastery. After judging by the expert, the objective is 34 questions. Then, the test was distributed to 34 students in grade 9^{th} that already learn about heat transfer topic as a test in google form, it divided into two parts (A and B). The next step after conducting the test, the result is analyzed by using ANATES software. Based on the analysis of ANATES, there were 15 number of test items used as the research instrument to measure students' concept mastery. According to the result of ANATES analysis, the reliability of the objective test was 0.79 which considered as high. The number of test items that were accepted as the research instrument to measure students' concept mastery can be seen in Table 3.8.

Table 3.8
Recapitulation Analysis Objective Test

Number of test item	Discrim -inant Power	Category	Difficulty Level	Category	Validity	Category	Status of test item	New number of test item
1	44.44	Good	55.88	Moderate	0.328	Low	Rejected	-
2	22.22	Satisfactory	88.24	Easy	0.378	Low	Revised	-
3	44.44	Good	76.47	Easy	0.371	Low	Revised	1
4	11.11	Poor	88.24	Easy	0.01	Very low	Rejected	-
5	55.56	Good	26.47	Difficult	0.448	Enough	Accepted	14
6	88.89	Excellent	44.12	Moderate	0.541	Enough	Accepted	3
7	22.22	Satisfactory	55.88	Moderate	0.295	Low	Rejected	-
8	44.44	Good	64.71	Moderate	0.358	Low	Revised	-
9	44.44	Good	44.12	Moderate	0.465	Enough	Accepted	-
10	33.33	Satisfactory	23.53	Difficult	0.239	Low	Rejected	-
11	66.67	Good	44.12	Moderate	0.454	Enough	Accepted	5
12	0.00	Poor	2.94	Difficult	0.068	Very Low	Rejected	-
13	33.33	Satisfactory	82.35	Easy	0.38	Low	Accepted	-
14	11.11	Poor	26.47	Difficult	0.179	Very Low	Rejected	-
15	66.67	Good	38.24	Moderate	0.478	Enough	Accepted	6
16	33.33	Satisfactory	32.35	Moderate	0.246	Low	Rejected	-
17	66.67	Good	58.82	Moderate	0.428	Enough	Accepted	7
18	44.44	Good	70.59	Easy	0.431	Enough	Rejected	-
19	66.67	Good	67.65	Moderate	0.537	Enough	Accepted	8
20	55.56	Good	67.65	Moderate	0.503	Enough	Accepted	10
21	55.56	Good	29.41	Difficult	0.385	Low	Revised	15
22	44.44	Good	61.76	Moderate	0.487	Enough	Accepted	12
23	44.44	Good	50.00	Moderate	0.437	Enough	Accepted	13
24	44.44	Good	76.47	Easy	0.409	Enough	Accepted	2
25	-22.22	Poor	44.12	Moderate	0.046	Very Low	Rejected	-
26	55.56	Good	64.71	Moderate	0.527	Enough	Accepted	11
27	66.67	Good	44.12	Moderate	0.584	Enough	Accepted	-
28	88.89	Excellent	67.65	Moderate	0.71	High	Accepted	4
29	11.11	Poor	26.47	Difficult	0.118	Very Low	Rejected	-
30	-22.22	Poor	47.06	Moderate	0.144	Very Low	Rejected	-
31	33.33	Satisfactory	41.18	Moderate	0.251	Low	Rejected	-
32	66.67	Good	55.88	Moderate	0.556	Enough	Accepted	9
33	22.22	Satisfactory	29.41	Difficult	0.231	Low	Rejected	-
34	22.22	Satisfactory	17.65	Difficult	0.200	Low	Rejected	-

3) Research Instrument Analysis Expert Judgments Results

Aside from students' validation, the research equipment was also evaluated by experts. The expert judgement was provided by two IPSE lectures, and one physics teacher from a private school which used the Cambridge Curriculum. From that Objective test and CPAM rubric are being evaluated. Here are the results of this research instrument's evaluation from judgements:

a. Avoid repeating questions.
b. Make the multiple choices consistent in term of its length of the text.

c. Delete the questions that has unclear instructions or multiple choices.
d. Delete the questions that aren't proper for junior high school cognitive level.
e. There are some questions that are adopted from IGCSE past paper and the level of that questions are quite difficult for 8th grade students.
f. The procedures on the worksheet are clearly written and can be followed by the students.
g. The researcher already prepared the instrument very well.

4) Creativity Rubric

CPAM is used to assess students' creative thinking abilities through group projects. Besemer and Treffinger (1981) developed the Creative Product Analysis Matrix (CPAM), which consists of Novelty, Resolution, Elaboration, and Synthesis. It is intended to aid in more comprehensive product monitoring and to concentrate judges' attention on key product qualities. The creativity rubric design is set out in the Table 3.9.

Table 3.9
The Creative Product Analysis Matrix (CPAM) Design

No	Creative Product Dimension	Criteria	Score		
			1	2	3
1	Novelty	Germinal			
		Original			
		Transformational			
2	Resolution	Adequate			
		Appropriate			
		Logical			
		Useful			
		Valuable			
3	Elaboration & Synthesis	Attractive			
		Complex			
		Elegant			
		Organic			
		Well-crafted			

Expressive

(Besemer & Treffinger, 1981)

Each score is interpreted as the level of student product creativity in this rubric. A score of 3 denotes "High", a score of 2 denotes "Medium", and a score of 1 denotes "Low". In contrast, another expert judgement suggests that the indications supplied in the form of a creative product assessment should be explained more clearly. The schematic indicator for CPAM is shown in Appendix A.2.

5) The Implementation of STEAM based learning

During a classroom observation, the implementation of STEAM based learning is assessed. An observation sheet is a non-test instrument used to examine the learning experiences throughout the implementation of STEAM based learning, whether each stages is executed or not. The observation sheet used in this research can be seen in Table 3.10.

Table 3.10
STEAM-PjBL Observation Sheet

STEAM-PjBL Stages	Activity	Implementation		%	Interpretation
		Yes	No		
Preparation	Triggering prior knowledge Dividing individual task Determining the material Creating the design project				
Implementation	Creating the project based on the design project Conducting the test of the product Identifying the problem of product				
Presentation	Presenting design idea and product Sharing problem				

	Sharing the solution to resolve the problem
Evaluation	Evaluating the project
Correction	Correcting the product based on the evaluation

Following the collection data, the data is measured using percentages in the observation sheet. Each stage's is examined by checking the "Yes" or "No" boxes in the observation sheet. Following that, the amount of "Yes" or "No" are summed up and converted into percentage. The formula for converting a score into percentage is as follows:

$$\%LI = \frac{\Sigma IA}{\Sigma TA} \times 100\%$$

Information:

LI = the implementation of learning experiences

IA = the amount of implemented aspect

TA = total amount of the aspect

(Ridwan, 2010, as cited in Salikha, 2020)

After the percentage of the implementation is obtained, it is interpreted into several categories, as shown in Table 3.11.

Table 3.11
The Interpretation of Learning Experiences Implementation

Percentage (%)	Interpretation
LI = 0	None of the learning experience is implemented
0 < LI < 25	A small number of learning experience are implemented
25 ≤ LI < 50	Almost half of the learning experiences are implemented
LI = 50	Half of the learning experiences are implemented
50 ≤ LI < 75	Most of the learning experiences are implemented

$75 \leq$ LI < 100	Almost all of the learning experiences are implemented
100	All of the learning experiences are implemented

(Ridwan, 2010, as cited in Salikha, 2020)

3.9 Data Analysis Technique

In this research, data was gathered using a quantitative technique. This methodology was used to analyze students' conceptual mastery and creativity. The following information is provided in detail below:

1) Students Concept Mastery

Students' concept mastery was obtained from 15 numbers of multiple-choice tests. The analysis was performed by using SPSS 25 software and Microsoft excels to calculate the percentage of the students who can answer the question correctly. The normality and homogeneity of the test items were examined using the software. The data collected were evaluated based on the level of cognitive domain which consisted of 3 levels: C1 (remembering), C2 (understanding), C4 (analyzing). The data from students' score is explained below.

a. Normalized Gain

To find the improvement of students' concept mastery score from pretest to posttest, normalized gain (N-gain) score is calculated. The N-gain was calculated using the formula below.

$$N - gain = \frac{Posttest\ score - pretest\ score}{Maximal\ score - pretest\ score}$$

(Hake, 1999)

After the N-gain score obtained, the result is categorized whether it is high, medium, or low. The interpretation of N-gain score can be seen in Table 3.12.

Table 3.12
N-gain Interpretation

N-gain Score	Category
N-gain > 0.7	High
0.7 ≥ N-gain ≥ 0.3	Medium
N-gain < 0.3	Low

(Hake, 1999)

b. Normality and Homogeneity Test

Normality and homogeneity test was conducted using SPSS 25 software. The normality test is done to see whether the data is normally distributed. Referencing to the significant value of Shapiro-Wilk, if the significance is ≥ 0.05, the data is normally distributed. Also, for the homogeneity test, if the significance value is ≥ 0.05, then the data is homogenous.

c. Wilcoxon Signed Rank Test

The next step after taking normality and homogeneity test is testing the hypothesis. Wilcoxon signed-rank test is used to determine whether the hypothesis is accepted or rejected. Referring to Wilcoxon S-R Test, if the value of significance two tailed less than 0.05 ($p \leq 0.05$), it can be said there are significant differences between two variables for the same subject, which rejected H_0 (i.e., hypothesis of no differences). Meanwhile, if the value of significance two tailed more than 0.05 ($p \geq 0.05$), H_0 is accepted. SPSS 25 software is used to find the significance results.

d. The Effect Size

The last stage is determine the magnitude of the results from Wilcoxon S-R Test analysis. Effect sizes can be used to express the significance of results (Lakens, 2013). Often, effect sizes are not provided through the SPSS software and must be calculated by hand (Woodrow, 2014). The effect size was calculated using formula below:

$$r = \frac{z}{\sqrt{N}}$$

r = effect size

z = z value

N = number of observations of two time points

The interpretation of score of approximate value of effect size using Cohen (1988) (as cited in Woodrow, 2014) guidelines, which shown in Table 3.13.

Table 3.13
The Interpretation of Effect Size (r) Value

Effect size (r)	Criteria
0.50 – 1	Large Effect
0.30 – 0.49	Medium Effect
0.10 – 0.29	Small Effect

(Cohen, 1988, as cited in Woodrow, 2014)

2) Students' creativity

Students' creativity was investigated based on their product results. Students 'creativity indicators were assessed by using rubric scoring under creativity dimension from Besemer and Treffiger (1981) and converted into percentage. According to Purwanto (2013) (as cited in Kurniaman, Oktari, & Pahrurazi, 2020) the formula is as follows:

$$NP = \frac{R}{SM} \times 100\%$$

Where:

NP = the percent value sought or expected

R = Raw score obtained

SM = Maximum score

The interpretation of score percentage of student creativity is categorized into a certain criteria, which shown on Table 3.14.

Table 3.14
The Interpretation of Students' Creativity

Percentage (%)	Criteria
86 – 100	Very Good
76 – 85	Good
60 – 75	Enough
55 – 56	Lack
<54	Very Lack

(Purwanto, 2013, as cited in Kurniaman et al., 2020)

3.10 Research Procedure

To ensure that the study is organized systematically, there are three major stages in the research: preparation, implementation, and completion.

1) Preparation Stage

This stage seeks to allow the author to examine all the variables in this research before doing the research itself. The following information is provided in further detail:

a. Analysis the research problem to be investigated in this research. Formulating research objective.
b. Conduct literature study about students' creativity, concept mastery, STEAM based learning, and heat transfer topic to enrich researcher knowledge. All literature come from reliable and trusted sources such as books, e-books, journals, and articles as well.
c. Perform the study of literature to guide in the development of the instruments used in this study, consisting of the syntax for STEAM based learning, and the instrument of students' creativity and concept mastery.
d. Design instrument comprising of creativity product analysis matrix (CPAM) rubric and test item.
e. Construct a lesson plan, presentations, platforms, and worksheets that are used as teaching material to assist in the implementation of STEAM based learning.
f. Validate the research instrument by expert judgment and students who have learned about related topic, in this case the topic is heat transfer.
g. Revise the instrument of research based on the advice of expert judgment and students' validation.

2) Implementation Stage

At this stage, the researcher begins to perform research to gather data for the research itself. The following is the comprehensive information:

a. Conduct pre-test for assessing students' prior knowledge about heat transfer topic.
b. Conduct the teaching-learning process using STEAM based learning. The STEAM implementation regulations are evaluated based on observations sheets relating to the stages of STEAM based learning which are: preparation, implementation, presentation, evaluation, and correction.
c. Assess students' creativity using creativity product analysis matrix (CPAM) rubric based on students' projects in STEAM based learning implementation.

d. Assess students' mastery concept by conducting post-test.
e. Collect the data to be analyzed and proceed at the next stage.

3) Completion Stage

This is the last stage of the research, which implies that the data has been collected and analyzed. This stage's specific information is as follows:

a. Process and analyze the data, based on the type of instrument used for each variable, and provide any further conclusions drawn from the data.
b. Arrange further discussions focused on interpretation of the results.
c. Create conclusion and recommendation based on result and discussion.
d. Completing the research paper.

The researcher designed a flow chart that is shown in Figure 3.1 to provide an overview of the procedure below:

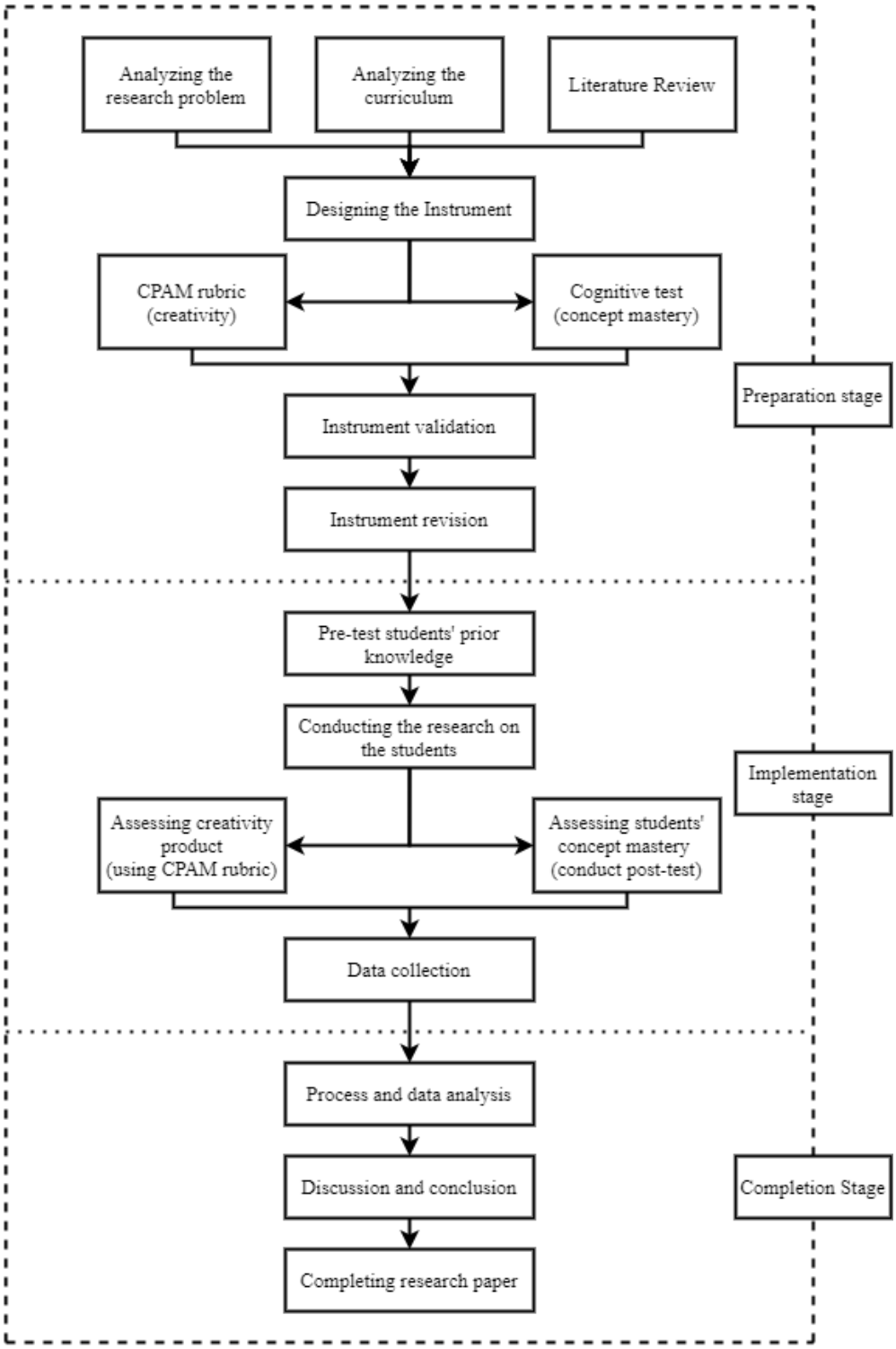

Figure 3.1 Flowchart of Research Procedures